

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 13, 2002, 10:08:42 ; Search time 23.07 Seconds  
(without alignments)  
120.467 Million cell updates/sec

Title: US-09-486-094-12

Perfect score: 51

Sequence: 1 XCXXXXXXCXXXXXXCX 19

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_17.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	56.9	117	10 Q40229	Q40229 lilium long
2	28	54.9	243	2 Q9R225	Q9R225 deinococcus
3	28	54.9	244	10 Q9SJD9	Q9SJD9 arabidopsis
4	28	54.9	454	10 Q9AU61	Q9AU61 triticum ae
5	28	54.9	462	10 Q39249	Q39249 arabidopsis
6	28	54.9	472	10 Q9SM43	Q9SM43 spinacia ol
7	28	54.9	473	10 Q40251	Q40251 lactuca sat
8	28	54.9	478	10 Q40593	Q40593 nicotiana t
9	28	54.9	776	2 Q47412	Q47412 escherichia
10	28	54.9	1297	5 Q9U350	Q9U350 caenorhabdi
11	28	54.9	2948	5 Q93785	Q93785 caenorhabdi
12	27	52.9	99	5 Q9N329	Q9N329 caenorhabdi
13	27	52.9	233	5 Q9NJ95	Q9NJ95 caenorhabdi
14	27	52.9	428	11 Q9D617	Q9D617 mus musculu
15	27	52.9	441	5 Q9W5X1	Q9W5X1 drosophila
16	27	52.9	565	10 Q9M195	Q9M195 arabidopsis
17	27	52.9	568	5 Q9NJ97	Q9NJ97 caenorhabdi
18	27	52.9	580	5 Q9VM37	Q9VM37 drosophila
19	27	52.9	2146	5 Q9VC97	Q9VC97 drosophila

20	27	52.9	4545	11 Q61291	Q61291 mus musculu
21	27	52.9	4599	4 Q9N2R2	Q9N2R2 homo sapien
22	26	51.0	58	5 Q9U623	Q9U623 pacifastacu
23	26	51.0	170	6 Q9TTZ7	Q9TTZ7 bos taurus
24	26	51.0	170	6 Q9TTZ6	Q9TTZ6 ovis aries
25	26	51.0	222	2 Q912L0	Q912L0 pseudomonas
26	26	51.0	246	4 Q99750	Q99750 homo sapien
27	26	51.0	246	11 Q99JM9	Q99JM9 mus musculu
28	26	51.0	248	5 Q96282	Q96282 plasmodium
29	26	51.0	318	10 Q41794	Q41794 zea mays (m
30	26	51.0	348	12 Q9J5S4	Q9J5S4 semliki for
31	26	51.0	438	12 Q87046	Q87046 semliki for
32	26	51.0	458	3 Q9UW11	Q9UW11 piromyces f
33	26	51.0	528	11 Q9CXD8	Q9CXD8 mus musculu
34	26	51.0	646	5 Q9XU22	Q9XU22 caenorhabdi
35	26	51.0	721	12 Q56281	Q56281 human herpe
36	26	51.0	724	10 Q9LW21	Q9LW21 arabidopsis
37	26	51.0	730	6 Q19051	Q19051 oryctolagus
38	26	51.0	731	6 Q19050	Q19050 oryctolagus
39	26	51.0	732	6 Q28484	Q28484 macaca fasc
40	26	51.0	997	11 Q9JJS0	Q9JJS0 mus musculu
41	26	51.0	999	4 Q9NQ36	Q9NQ36 homo sapien
42	26	51.0	1002	5 Q45247	Q45247 caenorhabdi
43	26	51.0	1145	12 Q87049	Q87049 semliki for
44	26	51.0	1581	13 Q73809	Q73809 fugu rubrip
45	26	51.0	4599	11 Q9J118	Q9J118 mus musculu

#### ALIGNMENTS

RESULT 1

Q40229 ID Q40229 PRELIMINARY: PRT; 117 AA.

AC Q40229; 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)

DE ORF (FRAGMENT).

GN LIM6.

OS Lilium longiflorum (Trumpet lily).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Lilium.

OX NCBI\_TaxID=4690;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. HINOMOTO; TISSUE=FLORAL BUD;

RA Tabata S.;

RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. HINOMOTO; TISSUE=FLORAL BUD;

RX MEDLINE=96051386; PubMed=7584025;

RA Kobayashi T., Kobayashi E., Sato S., Hotta Y., Miyazima N., Tanaka A., Tabata S.;

RA "Characterization of cDNAs induced in meiotic prophase in lily microsporocytes.";

RT DNA Res. 1:15-26(1994).

DR EMBL; D21812; BAA04836.1; -

DR Mendel; 12768; Lillo;1968;12768.

FT NON\_TER 1

SQ SEQUENCE 117 AA; 12625 MW; 972C71B1699CEA05 CRC64;

Query Match 56.9%; Score 29; DB 10; Length 117;  
Best Local Similarity 23.5%; Pred. No. 2.4;  
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2 CXXXXXXCXXXXXXC 18  
| | | | |  
Db 74 CTTSSCKKGGVTCCKC 90

```

RESULT 2
Q9RZ25 ID Q9RZ25 PRELIMINARY; PRT; 243 AA.
AC Q9RZ25;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE HYPOTHETICAL 26.2 KDA PROTEIN.
GN DRA0128.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001862; AAF12318.1; -
DR TIGR; DRA0128; -
DR InterPro; IPR001387; HTH_3.
DR InterPro; IPR000014; PAS.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
DR SMART; SM00091; PAS; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 243 AA; 26235 MW; E5105B3C1859FEFE CRC64;

Query Match 54.9%; Score 28; DB 2; Length 243;
Best Local Similarity 23.5%; Pred. No. 4.8;
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18
DB 214 CRWHSACRDTARCTSKC 230

RESULT 3
Q9SJD9 ID Q9SJD9 PRELIMINARY; PRT; 244 AA.
AC Q9SJD9;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE T27G7.23 (FRAGMENT).
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RA Kim C., Altafi H., Bei Q., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Liu J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T27G7 from chromosome
RT T.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006932; AAF22898.1; -
DR InterPro; IPR000566; Lipocin_cytFARP.

PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
NON_TER 244
SQ SEQUENCE 244 AA; 26631 MW; 4CD8221F444CF04C CRC64;

Query Match 54.9%; Score 28; DB 10; Length 244;
Best Local Similarity 23.5%; Pred. No. 4.8;
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18
DB 134 CIANPACAAVACLQTC 150

RESULT 4
Q9AU61 ID Q9AU61 PRELIMINARY; PRT; 454 AA.
AC Q9AU61;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE VIOLAXANTHIN DE-EPOXIDASE.
GN WVDE.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HANFENG 9703; TISSUE=LEAF;
RA Zou Q., Li B., Li Q.Z.;
RT "Cloning and gene expression of wheat violaxanthin de-epoxidase.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF265294; AAK38177.1; -
SQ SEQUENCE 454 AA; 51228 MW; 8165B5C6CF112774 CRC64;

Query Match 54.9%; Score 28; DB 10; Length 454;
Best Local Similarity 23.5%; Pred. No. 5.1;
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18
DB 127 CIANPSCAAVACLNTC 143

RESULT 5
Q39249 ID Q39249 PRELIMINARY; PRT; 462 AA.
AC Q39249;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE VIOLAXANTHIN DE-EPOXIDASE PRECURSOR.
GN AVDE1.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Bugos R.C., Yamamoto H.Y.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chip C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,

```

```

RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U44133; AAC50032.1; -.
DR EMBL; AC003981; AAF99753.1; -.
DR Mendel; 6341; Arath; Vdel; 6341.
DR InterPro; IPR000566; Lipocln_cytfabp.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 462 AA; 52017 MW; 58E37B2C12D4426B CRC64;

Query Match 54.9%; Score 28; DB 10; Length 462;
Best Local Similarity 23.5%; Pred. No. 5.1;
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18
Db 134 CIANPACAAVACLOTC 150

RESULT 6
Q9SM43
ID Q9SM43 PRELIMINARY; PRT; 472 AA.
AC Q9SM43;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE VIOLAXANTHIN DE-EPOXIDASE PRECURSOR.
GN SVDEL.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=LEAF;
RA Emanuelsson A.K., Eskling M., Akerlund H.E.;
RT "Cloning and sequencing of Spinacia oleracea violaxanthin de-
epoxidase."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ250433; CAB59211.1; -.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 472 AA; 53658 MW; 12B4C8C69AB61E12 CRC64;

Query Match 54.9%; Score 28; DB 10; Length 472;
Best Local Similarity 23.5%; Pred. No. 5.1;
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18
Db 145 CIANPSCAAVACLOTC 161

RESULT 7
Q40251
ID Q40251 PRELIMINARY; PRT; 473 AA.
AC Q40251;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VIOLAXANTHIN DE-EPOXIDASE PRECURSOR.
GN VDEL.
OS Lactuca sativa (Garden lettuce).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca.

```

```

OX NCBI_TaxID=4236;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ROMAINE;
RX MEDLINE=96270536; PubMed=8692813;
RA Bugos R.C., Yamamoto H.Y.;
RT "Molecular cloning of violaxanthin de-epoxidase from romaine lettuce
and expression in Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:6320-6325(1996).
DR EMBL; U31462; AAC49373.1; -.
DR Mendel; 8691; Lacsa; Vdel; 8691.
DR InterPro; IPR000566; Lipocln_cytfabp.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
KW Transit peptide.
FT TRANSIT.
FT CHAIN.
SQ SEQUENCE 473 AA; 54447 MW; 1B22522DC2C62699 CRC64;

Query Match 54.9%; Score 28; DB 10; Length 473;
Best Local Similarity 23.5%; Pred. No. 5.1;
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18
Db 146 CIANPSCAAVACLOTC 162

RESULT 8
Q40593
ID Q40593 PRELIMINARY; PRT; 478 AA.
AC Q40593;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE VIOLAXANTHIN DE-EPOXIDASE PRECURSOR.
GN TVDEL.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=XANTHI; TISSUE=LEAF;
RA Bugos R.C., Yamamoto H.Y.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U34817; AAC50031.1; -.
DR Mendel; 9222; Nicta; Vdel; 9222.
DR InterPro; IPR000566; Lipocln_cytfabp.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
KW Transit peptide.
FT TRANSIT.
FT CHAIN.
SQ SEQUENCE 478 AA; 54561 MW; 0967DF4547D7809D CRC64;

Query Match 54.9%; Score 28; DB 10; Length 478;
Best Local Similarity 23.5%; Pred. No. 5.1;
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18
Db 155 CIANPACAAVACLOTC 171

RESULT 9
Q47412
ID Q47412 PRELIMINARY; PRT; 776 AA.
AC Q47412;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

```

DE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE B1 SUBUNIT ALPHA-PRIME-  
DE POLYPEPTIDE PRECURSOR.

GN Escherichia coli.  
OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84272624; PubMed=6087316;  
RA Carlson J., Fuchs J.A., Messing J.;  
RT "Primary structure of the Escherichia coli ribonucleoside diphosphate  
reductase operon."  
RL Proc. Natl. Acad. Sci. U.S.A. 81:4294-4297(1984).  
DR EMBL; K02672; CAB33589.1; -;  
DR HSSP; P00452; 5RIR.  
DR InterPro; IPR00788; Ribonucleo\_red.  
DR Pfam; PF00317; ribonucleo\_red; 4.  
DR PRINTS; PR01183; RIBORDASEM1.  
FT CHAIN 27 776 POTENTIAL.  
FT CHAIN 3 776 POTENTIAL.  
SQ SEQUENCE 776 AA; 87626 MW; 152185957CEC4CE6 CRC64;

Query Match 54.9%; Score 28; DB 2; Length 776;  
Best Local Similarity 23.5%; Pred. No. 5.4;  
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18  
DB 654 CRTTCTCTPMSCGK 670

## RESULT 10

ID Q9U350 PRELIMINARY; PRT; 1297 AA.  
AC Q9U350;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE W05B2.2 PROTEIN.  
GN W05B2.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gardner A.E.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology."  
RT Science 282:2012-2018(1998).  
CC -1- SIMILARITY: TO PANCREATIC TRYPSIN INHIBITOR (KUNITZ) DOMAIN.  
DR EMBL; Z81138; CAB63320.1; -;  
DR HSSP; P31713; 1SHF.  
DR InterPro; IPR02899; EB.  
DR InterPro; IPR002223; Kunitz\_BPTI.  
DR Pfam; PF01683; EB; 2.  
DR Pfam; PF00014; Kunitz\_BPTI; 6.  
DR PRINTS; PR00759; BASICPTASE.  
DR SMART; SM00131; KU; 6;  
DR SMART; SM00289; WRI; 11.  
DR PROSITE; PS00280; BPTI\_KUNITZ\_1; 2.  
DR PROSITE; PS0279; BPTI\_KUNITZ\_2; 6.  
KW Serine protease inhibitor.  
SQ SEQUENCE 1297 AA; 143867 MW; 6599F3D479A3A7DD CRC64;

Query Match 54.9%; Score 28; DB 5; Length 1297;  
Best Local Similarity 23.5%; Pred. No. 5.8;  
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18  
DB 52 CTRNTECVSEATCLGTC 68

## RESULT 11

ID Q93785 PRELIMINARY; PRT; 2948 AA.  
AC Q93785;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE F54E4.1 PROTEIN.  
GN F54E4.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Thomas K.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans".  
RL Nature 368:32-38(1994).  
DR EMBL; Z79639; CAB01916.1; -;  
DR InterPro; IPR001680; WD40.  
DR Pfam; PF00400; WD40; 6.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR SMART; SM00320; WD40; 6.  
DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
DR PROSITE; PS00082; WD\_REPEATS\_2; 2.  
DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 2948 AA; 330689 MW; 199DC767D3B43448 CRC64;

Query Match 54.9%; Score 28; DB 5; Length 2948;  
Best Local Similarity 23.5%; Pred. No. 6.3;  
Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18  
DB 733 CTGLTVCNPSPPSCFISC 749

## RESULT 12

ID Q9N329 PRELIMINARY; PRT; 99 AA.  
AC Q9N329;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE HYPOTHETICAL PROTEIN Y59E9AL.5.  
GN Y59E9AL.5.  
OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).

RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=BRISTOL N2;

RA Waterston R.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC024838; AAF60819.1; -.

SQ SEQUENCE 99 AA; 11325 MW; EB3CC5BF97DAECC CRC64;

Query Match 52.9%; Score 27; DB 5; Length 99;

Best Local Similarity 23.5%; Pred. No. 7.8;

Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18

Db 78 CTSASRCSCSPCCSC 94

RESULT 13

Q9NJ95 ID Q9NJ95 PRELIMINARY; PRT; 233 AA.

AC Q9NJ95;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE NHR-25 (FRAGMENT).

GN NHR-25

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=N2;

RX MEDLINE=20237592; PubMed=10772806;

RA Gissendanner C.R., Sluder A.E.;

RT "nhr-25, the Caenorhabditis elegans ortholog of ftz-f1, is required

RT for epidermal and somatic gonad development.";

RL Dev. Biol. 221:259-272(2000).

CC -/- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -/- SIMILARITY: TO C4-TYPE STEROID RECEPTOR ZINC FINGER FAMILY.

DR EMBL; AF179216; AAF67040.1; -.

DR InterPro; IPR001628; zf-C4.

DR Pfam; PF00105; zf-C4; 1.

DR SMART; SM00047; STROIDFINGER.

DR SMART; SM00399; znf\_C4; 1.

DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.

DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;

KW zinc-finger.

FT NON\_TER 233 .. 233

SQ SEQUENCE 233 AA; 26601 MW; 8387EA2B35646625 CRC64;

Query Match

Best Local Similarity 52.9%; Score 27; DB 5; Length 233;

Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18

Db 54 CSAEANCHVDTRCRC 70

RESULT 14

Q9D6I7 ID Q9D6I7 PRELIMINARY; PRT; 428 AA.

AC Q9D6I7;

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE 2900024C23RIK PROTEIN.

GN 2900024C23RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=HIPPOCAMPUS;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK013580; BAB28914.1; -.

DR MGD; MGI:1914516; 2900024C23Rik.

SQ SEQUENCE 428 AA; 48936 MW; 7111D7C81F082822 CRC64;

Query Match

Best Local Similarity 52.9%; Score 27; DB 11; Length 428;

Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2 CXXXXXXCXXXXXXC 18

Db 331 CESDLDVYGTDCRTSC 347

RESULT 15

Q9W5X1 ID Q9W5X1 PRELIMINARY; PRT; 441 AA.

AC Q9W5X1;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE CG9572 PROTEIN.

GN CG9572

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

Wed Feb 13 17:11:44 2002

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., P.A.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE002611; AAF45369.1; -.  
 DR HSP; P15358; 1SKZ.  
 DR FlyBase; FBgn0031089; CG9572.  
 DR InterPro; IPR000561; EGF-like.  
 DR SMART; SM00181; EGF; 7.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN 4.  
 SQ SEQUENCE 441 AA; 46801 MW; 8021D0304127B97B CRC64;

Query Match 52.9%; Score 27; DB 5; Length 441;  
 Best Local Similarity 23.5%; Pred. No. 9.2;  
 Matches 4; Conservative 0; Mismatches 13; Indels 0; Gaps 0;  
 QY 2 CXXXXXXCXXXXXXC 18  
 DB 214 CSSRDMCSSNFKCVPAC 230

Search completed: February 13, 2002, 10:10:48  
 Job time: 126 sec